

REPORT DOCUMENTATION PAGE			Form Approved OMB NO. 0704-0188		
<p>The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA, 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p> <p>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</p>					
1. REPORT DATE (DD-MM-YYYY)		2. REPORT TYPE Technical Report		3. DATES COVERED (From - To) -	
4. TITLE AND SUBTITLE MODELING MULTIPLE RISKS: HIDDEN DOMAIN OF ATTRACTION			5a. CONTRACT NUMBER W911NF-10-1-0289		
			5b. GRANT NUMBER		
			5c. PROGRAM ELEMENT NUMBER 611102		
6. AUTHORS A. Mitra, S.I. Resnick			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAMES AND ADDRESSES Cornell University Office of Sponsored Programs Cornell University Ithaca, NY 14853 -2801			8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211			10. SPONSOR/MONITOR'S ACRONYM(S) ARO		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S) 57856-MA.24		
12. DISTRIBUTION AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.					
14. ABSTRACT Hidden regular variation is a sub-model of multivariate regular variation and facilitates accurate estimation of joint tail probabilities. We generalize the model of hidden regular variation to what we call hidden domain of attraction. We exhibit examples that illustrate the need for a more general model and discuss detection and estimation techniques.					
15. SUBJECT TERMS Regular variation, maximal domain of attraction, spectral measure, risk sets.					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Sidney Resnick
a. REPORT UU	b. ABSTRACT UU	c. THIS PAGE UU			19b. TELEPHONE NUMBER 607-255-1210

Report Title

MODELING MULTIPLE RISKS: HIDDEN DOMAIN OF ATTRACTION

ABSTRACT

Hidden regular variation is a sub-model of multivariate regular variation and facilitates accurate estimation of joint tail probabilities. We generalize the model of hidden regular variation to what we call hidden domain of attraction. We exhibit examples that illustrate the need for a more general model and discuss detection and estimation techniques.

MODELING MULTIPLE RISKS: HIDDEN DOMAIN OF ATTRACTION

ABHIMANYU MITRA AND SIDNEY I. RESNICK

ABSTRACT. Hidden regular variation is a sub-model of multivariate regular variation and facilitates accurate estimation of joint tail probabilities. We generalize hidden regular variation to what we call hidden domain of attraction. We exhibit examples that illustrate the need for a more general model and discuss detection and estimation techniques.

1. INTRODUCTION

Tail probabilities giving simultaneous exceedences of large thresholds by components of a risk vector provide risk measures for applications such as finance [21], environmental protection [25] and hydrology [3, 8]. Multivariate extreme value theory (MEVT) is a tool to approximate such tail probabilities but in common circumstances the asymptotic technique gives an incorrect tail probability approximation of 0. This paper points out that even when hidden regular variation (HRV) [11, 15, 16, 18, 20, 22, 23] is not applicable, a more general concept called *hidden domain of attraction* may yield a fix.

The joint distribution $H(\cdot)$ of a bivariate random vector $\mathbf{X} = (X^1, X^2)$ belongs to the maximal domain of attraction of a bivariate distribution $G(\cdot)$ if there exist scaling and centering constants $a_n^i > 0$ and $b_n^i \in \mathbb{R}$, $i = 1, 2$, such that for all continuity points $\mathbf{x} = (x^1, x^2)$ of G ,

$$(1.1) \quad \lim_{n \rightarrow \infty} [H(a_n^1 x^1 + b_n^1, a_n^2 x^2 + b_n^2)]^n = G(x^1, x^2)$$

and both the marginal distributions of $G(\cdot)$, $G^1(\cdot)$ and $G^2(\cdot)$, are non-degenerate extreme value distributions [9, page 208]. Let $M_+(\mathbb{E})$ be the set of Radon measures on \mathbb{E} and denote vague convergence by \xrightarrow{v} . The convergence (1.1) is equivalent to vague convergence in $M_+(\mathbb{E})$,

$$(1.2) \quad nP \left[\left(\frac{X^1 - b_n^1}{a_n^1}, \frac{X^2 - b_n^2}{a_n^2} \right) \in \cdot \right] \xrightarrow{v} \nu(\cdot) \quad (n \rightarrow \infty),$$

where depending on the case, \mathbb{E} is one of the following domains: $[-\infty, \infty]^2 \setminus \{(-\infty, -\infty)\}$ or $[0, \infty]^2 \setminus \{(0, 0)\}$ or $[-\infty, \infty] \times [0, \infty] \setminus \{(-\infty, 0)\}$ or $[0, \infty] \times [-\infty, \infty] \setminus \{(0, -\infty)\}$. The different domains correspond to the marginal limits being different extreme value distributions. The limit measure $\nu(\cdot)$ in (1.2) is related to the limit distribution $G(\cdot)$ in (1.1) by

$$(1.3) \quad \nu(\{(z^1, z^2) \in \mathbb{E} : z^1 \leq x^1, z^2 \leq x^2\}^c) = -\log(G(x^1, x^2)), \quad (\mathbf{x} = (x^1, x^2) \in \mathbb{E}).$$

Assuming $[X^1 > u, X^2 > v]$ is a rare event, that is, that u and v are sufficiently large, we use MEVT to approximate the joint tail probability $P(X^1 > u, X^2 > v)$ as

$$(1.4) \quad P(X^1 > u, X^2 > v) \approx \frac{1}{n} \nu \left(\left(\frac{u - b_n^1}{a_n^1}, \infty \right] \times \left(\frac{v - b_n^2}{a_n^2}, \infty \right] \right).$$

However, in the presence of asymptotic independence [9, page 226], (1.4) approximates this probability as zero. This approximation is often inaccurate and a better approximation is required.

Key words and phrases. Regular variation, maximal domain of attraction, spectral measure, risk sets.

S. Resnick and A. Mitra were partially supported by ARO Contract W911NF-10-1-0289 at Cornell University.

If (1.2) holds with $\mathbb{E} = [0, \infty]^2 \setminus \{(0, 0)\}$, $b_n^1 = b_n^2 = 0$ and some $a_n^1, a_n^2 \uparrow \infty$, then H possesses multivariate regular variation (MRV). If X^1 and X^2 are also asymptotically independent, we may improve the approximation of joint tail probabilities if hidden regular variation (HRV) is present; see [19, 20, 22, 23] and the seminal [15, 17]. However, HRV requires the distribution of $X^1 \wedge X^2$ to have a regularly varying tail and this may not be the case. Perhaps $X^1 \wedge X^2$ has a distribution in some maximal domain of attraction other than the heavy tailed domain. In this case, HRV cannot be applied to improve joint tail probability approximation but the deficiency can be remedied by a more general approach which we call *hidden domain of attraction* (HDA). HRV is a special case of HDA.

If the distribution of \mathbf{X} does not have MRV but (1.2) still holds, we may retrieve the MRV setup by transforming the components of \mathbf{X} to $(U^1(X^1), U^2(X^2))$, where $U^i(\cdot) = 1/(1 - H^i(\cdot))$ and $H^i(\cdot)$ is the distribution of X^i , $i = 1, 2$ [24, page 265]. If X^1 and X^2 are asymptotically independent, so are $(U^1(X^1), U^2(X^2))$ and assuming $U^1(X^1) \wedge U^2(X^2)$ has a regularly varying tail, we may seek HRV. Statistically this is problematic since we do not know $U^i(\cdot)$, $i = 1, 2$. This can be dealt with in various ways, none of which is completely satisfying or easy and a potential advantage in some circumstances of the notion of HDA is that it does not require that we transform components. Performing such transformations on data either introduces dependencies or unquantifiable errors, depending on the transformation method used. If $U^1(X^1) \wedge U^2(X^2)$ does not have a distribution with a regularly varying tail, HRV is not applicable but HDA may be if $U^1(X^1) \wedge U^2(X^2)$ is in a maximal domain of attraction of an extreme value distribution other than the Frechét. (See also Section 3.1.1.)

1.1. Outline. Section 1.2 reviews frequently used notation. In Section 2, we define hidden domain of attraction for the standard case, when both the components of the risk vector have the same distribution. This assumption is often unrealistic but could apply either when assessing serial tail dependence by sampling selectively from a stationary time series or if one chooses to standardize so that each marginal distribution is, say, Pareto. Section 3 deals with HDA in the non-standard case, where we drop the identical distribution assumption for X^1, X^2 . Having different marginals comes at the price that it is difficult to create a coherent estimation technique. We outline one estimation approach that works best when component random variables have the same upper bounds possibly ∞ . In both Sections 2 and 3, we exhibit examples which satisfy our model and discuss estimation procedures of limit measures that appear in the limit relations of the model. Section 4 discusses the detection techniques for HDA and estimation of joint and marginal tail probabilities. We conclude with a few remarks in Section 5.

1.2. Notation. For simplicity, this paper is restricted to two dimensions. For denoting a vector and its components, we use $\mathbf{x} = (x^1, x^2)$, $x^i = i$ -th component of \mathbf{x} , $i = 1, 2$. Multivariate intervals or rectangles are denoted $(\mathbf{x}, \mathbf{y}]$, $[\mathbf{x}, \mathbf{y}]$, etc where, for instance, $(\mathbf{x}, \mathbf{y}] = (x^1, y^1] \times (x^2, y^2]$. The vectors of all zeros, all ones and all infinities are denoted by $\mathbf{0} = (0, 0)$, $\mathbf{1} = (1, 1)$ and $\infty = (\infty, \infty)$ respectively. We write $x^{(1)} = x^1 \vee x^2$, $x^{(2)} = x^1 \wedge x^2$. So, the superscripts denote components of a vector and the ordered component is denoted by a parenthesis in the superscript.

We express vague convergence [23, page 173] of Radon measures as \xrightarrow{v} and weak convergence of probability measures [2, page 14] as \Rightarrow . Denote a point measure with points $\{x_i\}$ in a nice space \mathbb{F} by $\sum_i \epsilon_{x_i}$ where $\epsilon_x(\cdot)$ is the measure with all mass at x :

$$\epsilon_x(B) = \begin{cases} 1, & \text{if } x \in B, \\ 0, & \text{if } x \in B^c, \end{cases} \quad (x \in \mathbb{F}, B \subset \mathbb{F}).$$

The set of non-negative Radon measures on a space \mathbb{F} topologized by the vague topology is $M_+(\mathbb{F})$.

For a one dimensional distribution $F(x)$, set $\bar{F} := 1 - F$. The inverse of a non-decreasing function $\psi(x)$ is $\psi^\leftarrow(x) := \inf\{y : \psi(y) \geq x\}$. The space of left continuous functions on $(0, \infty]$ with finite right limits is $D_{\text{left}}((0, \infty])$ and the right continuous functions with finite left limits is $D((0, \infty])$.

2. STANDARD CASE HIDDEN DOMAIN OF ATTRACTION

Suppose that a bivariate random vector $\mathbf{X} = (X^1, X^2)$ with distribution $H(\mathbf{x})$ belongs to the maximal domain of attraction of an extreme value distribution G [24, page 265], $X^1 \stackrel{d}{=} X^2$ and X^1 and X^2 are asymptotically independent so that (1.1) is satisfied with $b_n^1 = b_n^2 = b_n$, $a_n^1 = a_n^2 = a_n > 0$ and $G(x, y) = G^1(x)G^2(y)$ for $x, y \in \mathbb{R}$, $n \in \mathbb{N}$. Thus, (1.2) becomes

$$(2.1) \quad nP[(\mathbf{X} - b_n \mathbf{1})/a_n \in \cdot] \xrightarrow{v} \nu(\cdot)$$

in $M_+(\mathbb{E})$, where $\mathbb{E} = [-\infty, \infty] \setminus \{-\infty\}$ or $\mathbb{E} = [0, \infty] \setminus \{0\}$. Since $G(x, y) = G^1(x)G^2(y)$ for $x, y \in \mathbb{R}$, the relation of ν and G given in (1.3) gives for $\mathbf{x} \in \mathbb{E}$,

$$(2.2) \quad \begin{aligned} \nu(\{\mathbf{z} \in \mathbb{E} : z^1 \leq x^1, z^2 \leq x^2\}^c) &= -\log G^1(x^1) + -\log G^2(x^2) \\ &= \nu(\{\mathbf{z} \in \mathbb{E} : z^1 \leq x^1\}^c) + \nu(\{\mathbf{z} \in \mathbb{E} : z^2 \leq x^2\}^c). \end{aligned}$$

The standard case contains the additional assumption that $X^1 \stackrel{d}{=} X^2$, which reduces (1.2) to (2.1) and reduces possible choices for \mathbb{E} . The cone $\mathbb{E} = [0, \infty] \setminus \{0\}$ is chosen only when H has MRV.

From (2.1), the maximal component of \mathbf{X} satisfies as $n \rightarrow \infty$,

$$(2.3) \quad nP(X^{(1)} > a_n y + b_n) \rightarrow \nu(\{\mathbf{z} \in \mathbb{E} : z^{(1)} > y\}), \quad (y, y) \in \mathbb{E},$$

so $X^{(1)}$ is in a maximal domain of attraction of an extreme value distribution and the distribution of $X^{(1)}$ characterizes a_n and b_n given in (2.1). Using one-dimensional extreme value theory, we can and do choose a_n and b_n in such a way that

$$(2.4) \quad \psi(y) := (\nu(\{\mathbf{z} \in \mathbb{E} : z^{(1)} > y\}))^{-1}$$

takes one of the following forms:

$$\psi(y) = \begin{cases} y^{1/\gamma}, & \text{if } y > 0, \\ 0, & \text{otherwise,} \end{cases} \quad \text{if } \gamma > 0, \\ \psi(y) = e^y, & \text{if } \gamma = 0, \\ \psi(y) = \begin{cases} \infty, & \text{if } y > 0, \\ (-y)^{1/\gamma}, & \text{otherwise,} \end{cases} \quad \text{if } \gamma < 0,$$

where γ is the extreme value index of the distribution of $X^{(1)}$ [24, page 9].

We define a sub-model of MEVT called (standard case) hidden domain of attraction (HDA). HDA helps approximate joint tail probabilities in the presence of asymptotic independence and includes HRV as a special case. We assume \mathbb{E} is either $[-\infty, \infty] \setminus \{-\infty\}$ or $[0, \infty] \setminus \{0\}$, and define \mathbb{E}^0 as either $(-\infty, \infty]$ or $(0, \infty]$. It is not always true that $\mathbb{E}^0 \subset \mathbb{E}$ (see Remark 2.2(2)) and we may have $\mathbb{E}^0 = (-\infty, \infty]$ and $\mathbb{E} = [0, \infty] \setminus \{0\}$. See Example 2.7. The choice of \mathbb{E}^0 depends on the domain of attraction of the distribution of $X^1 \wedge X^2$ and the choice of \mathbb{E} depends on the domain of $X^1 \vee X^2$.

Definition 2.1. The distribution of $\mathbf{X} = (X^1, X^2)$ has standard case hidden domain of attraction on the cone \mathbb{E}^0 if (i) $X^1 \stackrel{d}{=} X^2$; (ii) (2.1) and (2.2) hold; (iii) there exist positive scaling and real

centering constants $\{c_n\}$ and $\{d_n\}$ and a non-zero measure $\nu^0 \in M_+(\mathbb{E}^0)$ such that in $M_+(\mathbb{E}^0)$,

$$(2.5) \quad nP[(\mathbf{X} - d_n \mathbf{1})/c_n \in \cdot] \xrightarrow{v} \nu^0(\cdot) \quad (n \rightarrow \infty).$$

This definition requires $X^1 \stackrel{d}{=} X^2$ and that the distribution of \mathbf{X} belongs to the maximal domain of attraction of an extreme value product measure G with exponent measure ν . Some other remarks:

- Remark 2.2.** (1) HRV assumes (2.1) holds on $\mathbb{E} = [\mathbf{0}, \infty] \setminus \{\mathbf{0}\}$ with $b_n = 0$, $a_n \uparrow \infty$ and (2.5) is satisfied on $\mathbb{E}^0 = (\mathbf{0}, \infty]$ with $d_n = 0$, $c_n \uparrow \infty$ with $a_n/c_n \rightarrow \infty$ as $n \rightarrow \infty$. Hidden regular variation is a special case of hidden domain of attraction. HRV is the only sub-model of HDA where the cone \mathbb{E}^0 in (2.5) is $\mathbb{E}^0 = (\mathbf{0}, \infty]$.
 (2) From (2.5) the minimum component of \mathbf{X} satisfies,

$$(2.6) \quad nP[X^{(2)} > c_n y + d_n] \rightarrow \nu^0((y, \infty] \times (y, \infty]) \quad ((y, y) \in \mathbb{E}^0),$$

and therefore, the distribution of $X^{(2)}$ belongs to the maximal domain of attraction of an extreme value distribution [9, page 4]. When HRV exists, the distribution of $X^{(2)}$ has a regularly varying tail and is hence in the domain of attraction of the Fréchet distribution. HDA allows the additional cases where the distribution of $X^{(2)}$ belongs to the domain of attraction of the Gumbel or the Weibull distribution.

The distribution of $X^{(2)}$ determines the scaling and centering constants $\{c_n\}$ and $\{d_n\}$ and the cone \mathbb{E}^0 . As illustrated by Example 2.7, even if $\mathbb{E} = [\mathbf{0}, \infty] \setminus \{\mathbf{0}\}$, the cone \mathbb{E}^0 could be $(-\infty, \infty]$ and \mathbb{E}^0 is not necessarily a sub-cone of \mathbb{E} , as was the case for HRV [22].

- (3) From (2.3), $X^{(1)}$ belongs to the maximal domain of attraction of an extreme value distribution. Since $X^{(1)} \geq X^{(2)}$, $\gamma \geq \gamma^0$ and the convergence (2.1) puts a restriction on what convergences are possible in (2.5). For example, if (2.1) is satisfied with $X^{(1)}$ being in the Gumbel domain of attraction, then HRV can never hold on \mathbb{E}^0 since the tail of $X^{(2)}$ cannot be heavier than the tail of $X^{(1)}$.

2.1. Semi-parametric structure of ν^0 . The limit measure ν^0 in (2.5) has a semi-parametric structure which characterizes the limits in (2.5) as a class indexed by a real parameter and a set of probability measures. This semi-parametric form also assists estimation (as in [19, 20] for HRV).

To understand this semi-parametric structure, let $H^{(2)}(\cdot)$ be the distribution of $X^{(2)}$ and define

$$(2.7) \quad \psi^0(y) := [\nu^0((y, \infty] \times (y, \infty))]^{-1}$$

where $\nu^0(\cdot)$ is given in (2.5). Rewrite (2.6) as

$$(2.8) \quad \overline{nH^{(2)}}(c_n y + d_n) \rightarrow [\psi^0(y)]^{-1} \quad (y \in \mathbb{R}),$$

and univariate extreme value theory implies $H^{(2)}$ is in a maximal domain of attraction with some extreme value index γ^0 . Choose c_n and d_n suitably [24, page 9] and $\psi^0(\cdot)$ takes one of the forms:

$$(2.9) \quad \begin{aligned} \psi^0(y) &= \begin{cases} y^{1/\gamma^0}, & \text{if } y > 0, \\ 0, & \text{otherwise,} \end{cases} & \text{if } \gamma^0 > 0, \\ \psi^0(y) &= e^y, & \text{if } \gamma^0 = 0, \\ \psi^0(y) &= \begin{cases} \infty, & \text{if } y > 0, \\ (-y)^{1/\gamma^0}, & \text{otherwise,} \end{cases} & \text{if } \gamma^0 < 0. \end{aligned}$$

Henceforth assume that c_n and d_n are chosen so that (2.9) is true. Define

$$(2.10) \quad U^{(2)}(x) = 1/(1 - H^{(2)}(x)),$$

so that (2.8) becomes

$$(2.11) \quad U^{(2)}(c_n y + d_n)/n \rightarrow \psi^0(y), \quad y \in \mathbb{R}.$$

The following identifies the semi-parametric structure of ν^0 .

Proposition 2.3. *If $X^1 \stackrel{d}{=} X^2$, the distribution of \mathbf{X} satisfies (2.5) iff*

- (1) $X^{(2)}$ has a distribution $H^{(2)}$ in a univariate maximal domain of attraction so that for some $c_n > 0$, $d_n \in \mathbb{R}$, (2.8) or (2.11) holds, and
- (2) regular variation on $(\mathbf{0}, \infty]$ holds for the distribution of $(U^{(2)}(X^1), U^{(2)}(X^2))$:

$$(2.12) \quad nP \left[n^{-1} \left(U^{(2)}(X^1), U^{(2)}(X^2) \right) \in \cdot \right] \xrightarrow{v} \tilde{\nu}^0(\cdot)$$

where $U^{(2)}(\cdot)$ is defined in (2.10) and $\tilde{\nu}^0(\cdot)$ is a Radon measure on $(\mathbf{0}, \infty]$ that is related to the limit measure $\nu^0(\cdot)$ in (2.5) by

$$(2.13) \quad \tilde{\nu}^0((x^1, \infty] \times (x^2, \infty]) = \nu^0 \left(((\psi^0)^\leftarrow(x^1), \infty] \times ((\psi^0)^\leftarrow(x^2), \infty] \right), \quad (\mathbf{x} \in (\mathbf{0}, \infty]).$$

The measure $\tilde{\nu}^0(\cdot)$ satisfies the scaling property on $(\mathbf{0}, \infty]$:

$$(2.14) \quad \tilde{\nu}^0(c \cdot) = c^{-1} \tilde{\nu}^0(\cdot), \quad c > 0.$$

Remark 2.4. (i) Proposition 2.5 below shows that the limit measure $\tilde{\nu}^0$ is determined by a probability measure S^0 on a certain space $\delta\mathbb{N}^{(2)}$ to be explained and the family of limits in (2.12) is indexed by probability measures on $\delta\mathbb{N}^{(2)}$. Given the measure S^0 and the parameter γ^0 we get ν^0 which shows the semi-parametric structure. The probability measure S^0 determines $\tilde{\nu}^0$ and γ^0 gives $\psi^0(\cdot)$ from (2.9) and then ν^0 .

- (ii) If the support of the distribution of $X^{(2)}$ is smaller than that of X^i , $i = 1, 2$, then $U^{(2)}(X^i)$ could take the value ∞ with positive probability.

Proof of Proposition 2.3. Assume (2.8) or (2.11) as well as (2.12). For $\mathbf{x} \in \mathbb{E}^0$ such that $\psi^0(x^i) < \infty$, $i = 1, 2$, we have $c_n x^i + d_n$ converging from below to the right end point of $H^{(2)}$ and

$$\begin{aligned} nP \left[\frac{X^1 - d_n}{c_n} > x^1, \frac{X^2 - d_n}{c_n} > x^2 \right] \\ = nP \left[\frac{U^{(2)}(X^1)}{n} > \frac{U^{(2)}(c_n x^1 + d_n)}{n}, \frac{U^{(2)}(X^2)}{n} > \frac{U^{(2)}(c_n x^2 + d_n)}{n} \right] \\ \rightarrow \tilde{\nu}^0 \left((\psi^0(x^1), \infty] \times (\psi^0(x^2), \infty] \right) = \nu^0((x^1, \infty] \times (x^2, \infty]), \end{aligned}$$

where the convergence follows from (2.8) and (2.12) and the last equality follows from (2.13) and the forms of $\psi^0(\cdot)$ given in (2.9). Hence, (2.5) holds. The converse is similar and is omitted. \square

The scaling property (2.14) allows us to express (2.12) in an alternate coordinate system that transforms the limit measure into a product. From (2.12), $(U^{(2)}(X^1), U^{(2)}(X^2))$ has regular variation on $(\mathbf{0}, \infty]$ and using (2.7), (2.9) and (2.13), we get

$$\tilde{\nu}^0((\mathbf{1}, \infty]) = \nu^0 \left(((\psi^0)^\leftarrow(1), \infty]^2 \right) = [\psi^0((\psi^0)^\leftarrow(1))]^{-1} = 1.$$

The scaling property (2.14) extends this to $\tilde{\nu}^0([1, \infty]) = 1$ and Proposition 3.1 of [20] and Proposition 2.3 yield the equivalent convergence in alternate coordinates given in (2.15) below.

To specify the coordinates, we need the following: Let ν_1 be a Pareto measure on $(0, \infty]$ satisfying $\nu_1((y, \infty]) = y^{-1}$ for $y > 0$. Since $U^{(2)}(\cdot)$ is non-decreasing, $U^{(2)}(X^{(2)}) = U^{(2)}(X^1) \wedge U^{(2)}(X^2)$. Set

$$\delta\aleph^{(2)} = \{\mathbf{x} \in (0, \infty]^2 : x^{(2)} = 1\}.$$

Proposition 2.5. *The convergence in (2.12) is equivalent to*

$$(2.15) \quad nP \left[\left(\frac{U^{(2)}(X^{(2)})}{n}, \left(\frac{U^{(2)}(X^1)}{U^{(2)}(X^{(2)})}, \frac{U^{(2)}(X^2)}{U^{(2)}(X^{(2)})} \right) \right) \in \cdot \right] \xrightarrow{v} \nu_1 \times S^0(\cdot),$$

on $(0, \infty] \times \delta\aleph^{(2)}$, where S^0 is a probability measure on $\delta\aleph^{(2)}$. The relation between $\tilde{\nu}^0$ in (2.12) and S^0 is

$$(2.16) \quad \tilde{\nu}^0 \left(\left\{ \mathbf{x} \in (0, \infty]^2 : x^{(2)} \geq r, \mathbf{x}/x^{(2)} \in \Lambda \right\} \right) = r^{-1} S^0(\Lambda),$$

for $r > 0$ and Borel sets $\Lambda \subset \delta\aleph^{(2)}$.

The probability measure S^0 , called the *standardized hidden spectral measure*, is

$$(2.17) \quad P \left[\frac{(U^{(2)}(X^1), U^{(2)}(X^2))}{U^{(2)}(X^{(2)})} \in \cdot \mid X^{(2)} > t \right] \Rightarrow S^0(\cdot),$$

where the convergence holds as $t \rightarrow x_{H^{(2)}} = \sup\{y \in \mathbb{R} : H^{(2)}(y) < 1\}$. For HRV, a similar spectral measure was defined in [20]. In classical extreme value theory, a change to Pareto scale in the coordinates and then a change to polar coordinates produces a product measure and spectral measure ([9, page 214], [24, Chapter 5], [23, Chapter 6]).

2.2. Examples. We give examples of distributions that possess multivariate regular variation with asymptotic independence. Each has hidden domain of attraction but *not* hidden regular variation, emphasizing the need for a concept beyond HRV.

Example 2.6. Suppose, $W_1, W_2 \stackrel{iid}{\sim} F(\cdot)$ and $Z_1, Z_2 \stackrel{iid}{\sim} D(\cdot)$, where F and D belong to the maximal domains of attraction of the Fréchet with index $\alpha = 1$ and Gumbel distributions respectively. Let B be a Bernoulli random variable such that $P[B = 1] = 0.5 = 1 - P[B = 0]$. Assume the random variables W_1, W_2, Z_1, Z_2 and B are mutually independent and define a bivariate random vector \mathbf{X} as

$$\mathbf{X} = (X^1, X^2) = B(W_1, Z_1) + (1 - B)(Z_2, W_2).$$

We show that the distribution of \mathbf{X} has MRV. It suffices [7] to verify that $t_1 X^1 \vee t_2 X^2$ has a regularly varying tail for any $t_1, t_2 > 0$. Since F has a regularly varying tail and

$$P[t_1 W_1 \vee t_2 Z_1 > x] \sim P[t_1 W_1 > x], \quad (x \rightarrow \infty),$$

$t_1 X^1 \vee t_2 X^2$ also has a regularly varying tail. Thus for appropriate $a_n \uparrow \infty$,

$$nP[\mathbf{X}/a_n \in \cdot] \xrightarrow{v} \nu(\cdot)$$

on $\mathbb{E} = [0, \infty]^2 \setminus \{\mathbf{0}\}$, where $\nu([0, x^1] \times [0, x^2])^c = \frac{1}{2} \left((x^1)^{-1} + (x^2)^{-1} \right)$. Therefore, the distribution of \mathbf{X} has MRV with asymptotic independence on \mathbb{E} .

Furthermore, $X^{(2)}$ belongs to the maximal domain of attraction of a Gumbel distribution and therefore, HRV does not exist. To see this, without loss of generality [1, 24], assume that D is a von-Mises function [24, page 40] and for specificity assume the right endpoint of D is infinite. The form of the tail is

$$\overline{D}(x) = ce^{-\int_1^x [f_D(y)]^{-1} dy} \quad \text{and} \quad f_D'(x) \xrightarrow{x \rightarrow \infty} 0,$$

where $c > 0$ is some constant. Likewise, assume without loss of generality [24, page 58] that F satisfies $xF'(x)/\bar{F}(x) \rightarrow 1$, where 1 is the index of regular variation of \bar{F} . Then $\bar{H}^{(2)}(x) := \bar{F}(x)\bar{D}(x)$ is the tail of a von-Mises function with auxiliary function

$$f_{H^{(2)}}(t) = tf_D(t)/(t+1 \cdot f_D(t)) \stackrel{t \rightarrow \infty}{\sim} f_D(t),$$

since

$$\frac{1}{f_{H^{(2)}}(t)} = \frac{H^{(2)'}(t)}{H^{(2)}(t)} = \frac{F'(t)}{\bar{F}(t)} + \frac{D'(t)}{\bar{D}(t)} \stackrel{t \rightarrow \infty}{\sim} t^{-1} + \frac{1}{f_D(t)}.$$

From [24, Corollary 1.7, page 46] $H^{(2)}$ belongs to the maximal domain of attraction of the Gumbel distribution.

Next we make the conventional choices [24, page 40] of scaling and centering constants $\{c_n\}$ and $\{d_n\}$ in (2.8) so that in (2.9) $\psi^0(\cdot) = e^y$. These choices are $d_n = (1/H^{(2)})^{\leftarrow}(n)$ and $c_n = f_{H^{(2)}}(d_n)$. Then for $\mathbf{x} \in \mathbb{E}^0 = (-\infty, \infty]^2$, as $n \rightarrow \infty$,

$$\begin{aligned} nP[X^1 > d_n + c_n x^1, X^2 > d_n + c_n x^2] \\ &= \frac{n}{2} \bar{F}(d_n + c_n x^1) \bar{D}(d_n + c_n x^2) + \frac{n}{2} \bar{F}(d_n + c_n x^2) \bar{D}(d_n + c_n x^1) \\ &= \frac{n}{2} \left(\frac{\bar{F}(d_n + c_n x^1)}{\bar{F}(d_n + c_n x^2)} \right) \bar{H}^{(2)}(d_n + c_n x^2) + \frac{n}{2} \left(\frac{\bar{F}(d_n + c_n x^2)}{\bar{F}(d_n + c_n x^1)} \right) \bar{H}^{(2)}(d_n + c_n x^1) \\ &\rightarrow \frac{1}{2}(e^{-x^1} + e^{-x^2}). \end{aligned}$$

The convergence follows from the facts that \bar{F} is regularly varying, $c_n/d_n \rightarrow 0$ and (2.8) holds with $\psi^0(y) = e^y$. Therefore, as in Definition 2.1, the distribution of \mathbf{X} has HDA on $\mathbb{E}^0 = (-\infty, \infty]^2$ with limit measure ν^0 such that for $\mathbf{x} = (x^1, x^2) \in \mathbb{E}^0$,

$$\nu^0((x^1, \infty] \times (x^2, \infty]) = \frac{1}{2}(e^{-x^1} + e^{-x^2}).$$

To summarize, the distribution of \mathbf{X} is regularly varying on \mathbb{E} , has HDA on \mathbb{E}^0 , but does not have HRV since HRV requires the distribution of $X^{(2)}$ to be in the domain of attraction of the Fréchet distribution [24, page 54].

Example 2.7. Suppose, $U \sim \text{Uniform}([0, 1])$. Define the random vector \mathbf{X} as

$$\mathbf{X} = (X^1, X^2) = \left(\frac{1}{U}, \frac{1}{1-U} \right).$$

Now, note that for $x^1, x^2 > 0$, $2n > (x^1)^{-1} + (x^2)^{-1}$,

$$\begin{aligned} n(1 - P[1/U \leq 2nx^1, 1/(1-U) \leq 2nx^2]) &= n(1 - P[U \geq (2nx^1)^{-1}, U \leq 1 - (2nx^2)^{-1}]) \\ &= n(1 - (1 - (2nx^2)^{-1} - (2nx^1)^{-1})) \rightarrow \frac{1}{2}((x^1)^{-1} + (x^2)^{-1}), \end{aligned}$$

as $n \rightarrow \infty$. Therefore, on $\mathbb{E} = [0, \infty]^2 \setminus \{\mathbf{0}\}$, $nP[\mathbf{X}/2n \in \cdot] \xrightarrow{v} \nu(\cdot)$ where the limit measure ν satisfies $\nu([0, x^1] \times [0, x^2])^c = ((x^1)^{-1} + (x^2)^{-1})/2$, for $\mathbf{x} \in \mathbb{E}$ and thus the distribution of \mathbf{X} has MRV with asymptotic independence.

Since $X^1 \wedge X^2$ is bounded above, it is immediately clear HRV is absent. However, this distribution does have HDA. For $\{(x^1, x^2) \in (-\infty, \infty]^2 : x^1 + x^2 \leq 0\}$, and large n ,

$$\begin{aligned} nP\left[X^1 > 2 + \frac{2x^1}{n+1}, X^2 > 2 + \frac{2x^2}{n+1}\right] &= nP\left[U < \frac{n+1}{2n+2+2x^1}, U > \frac{n+1+2x^2}{2n+2+2x^2}\right] \\ &= n\left(\frac{n+1}{2n+2+2x^1} - \frac{n+1+2x^2}{2n+2+2x^2}\right) = n\left(\frac{1}{2} - \frac{x^1}{2n+2+2x^1} - \frac{1}{2} - \frac{x^2}{2n+2+2x^2}\right) \\ (2.18) \quad &\rightarrow \frac{1}{2}((-x^1) + (-x^2)), \end{aligned}$$

as $n \rightarrow \infty$. Similar calculations show that for $\{(x^1, x^2) \in (-\infty, \infty]^2 : x^1 + x^2 > 0\}$,

$$(2.19) \quad nP\left[X^1 > 2 + \frac{2x^1}{n+1}, X^2 > 2 + \frac{2x^2}{n+1}\right] \rightarrow 0,$$

as $n \rightarrow \infty$. Therefore, the distribution of \mathbf{X} has HDA as in Definition 2.1 on $\mathbb{E}^0 = (-\infty, \infty]^2$ with limit measure ν^0 such that for $\mathbf{x} \in \mathbb{E}^0$,

$$\nu^0((x^1, \infty] \times (x^2, \infty]) = \begin{cases} \frac{1}{2}((-x^1) + (-x^2)), & \text{if } x^1 + x^2 \leq 0 \\ 0 & \text{otherwise.} \end{cases}$$

From (2.18) and (2.19) it also follows that $X^{(2)}$ belongs to the domain of attraction of the reversed Weibull distribution [24, page 59].

To summarize: The distribution of \mathbf{X} has MRV with asymptotic independence, does not have HRV but does have HDA and furthermore, \mathbb{E}^0 is not a subset of \mathbb{E} .

2.3. Estimation. Recall the standard case assumes marginal distributions are the same. To estimate joint tail probabilities, we first estimate the limit measure ν^0 given in (2.5). Let, $\{\mathbf{X}, \mathbf{X}_i, i = 1, 2, \dots, n\}$ be iid with a common distribution satisfying (2.5). From (2.5),

$$(2.20) \quad \frac{1}{k} \sum_{i=1}^n \epsilon\left(\frac{\mathbf{X}_i - d(n/k)\mathbf{1}}{c(n/k)}\right)(\cdot) \Rightarrow \nu^0(\cdot) \quad (k \rightarrow \infty, n/k \rightarrow \infty),$$

in $M_+(\mathbb{E}^0)$ [23, page 139]. From (2.6), the distribution of $X^{(2)}$ determines c_n and d_n . The iid data $\{X_i^{(2)} : i = 1, 2, \dots, n\}$ allow estimates ([9], [23, page 93]) of $c(n/k)$ and $d(n/k)$, denoted by $\hat{c}(n/k)$ and $\hat{d}(n/k)$, satisfying

$$(2.21) \quad \frac{c(n/k)}{\hat{c}(n/k)} \xrightarrow{P} 1, \quad \frac{d(n/k) - \hat{d}(n/k)}{c(n/k)} \xrightarrow{P} 0;$$

Therefore, we get the joint convergence

$$(2.22) \quad \left(\frac{1}{k} \sum_{i=1}^n \epsilon\left(\frac{\mathbf{X}_i - d(n/k)\mathbf{1}}{c(n/k)}\right), \frac{d(n/k) - \hat{d}(n/k)}{c(n/k)}, \frac{c(n/k)}{\hat{c}(n/k)}\right) \Rightarrow (\nu^0(\cdot), 0, 1)$$

in $M_+(\mathbb{E}^0) \times \mathbb{R}^2$. Apply the almost surely continuous map $(\nu(\cdot), b, a) \mapsto \nu(a[(\cdot) + b\mathbf{1}])$ in (2.22) and we get the following proposition:

Proposition 2.8. *Let, $\{\mathbf{X}, \mathbf{X}_i, i \geq 1\}$ be iid with common distribution satisfying (2.5). Then,*

$$(2.23) \quad \widehat{\nu}_n^0(\cdot) := \frac{1}{k} \sum_{i=1}^n \epsilon\left(\frac{\mathbf{X}_i - \hat{d}(n/k)\mathbf{1}}{\hat{c}(n/k)}\right)(\cdot) \Rightarrow \nu^0(\cdot) \quad (k \rightarrow \infty, n/k \rightarrow \infty),$$

in $M_+(\mathbb{E}^0)$.

Estimation of $\nu^0(\cdot)$ in Proposition 2.8 does not exploit the semi-parametric structure discussed in Section 2.1 and has the disadvantages that (a) there is no guarantee the estimator $\widehat{\nu}_n^0(\cdot)$ is even a member of the class of possible limit measures; and (b) we are required to estimate $c(\cdot)$ and $d(\cdot)$. These problems are overcome using the semi-parametric structure as was done for HRV [11, 19, 20]. We must estimate the extreme value index γ^0 of the distribution of $X^{(2)}$ as well as the standardized hidden spectral measure S^0 . Since $\{X_i^{(2)} : i = 1, 2, \dots, n\}$ is iid data, estimating γ^0 of $X^{(2)}$ is a standard procedure [9, page 65] so we concentrate on estimating $S^0(\cdot)$. A modification of a ranks method [9, 11, 14, 23] to obtain an estimator of $\tilde{\nu}^0(\cdot)$ avoids the need to estimate $c(\cdot)$ and $d(\cdot)$. For $i = 1, 2, \dots, n$, define

$$(2.24) \quad R_i^{1,(2)} = |\{j : X_j^{(2)} \geq X_i^{(2)}\}| \quad \text{and} \quad R_i^{2,(2)} = |\{j : X_j^{(2)} \geq X_i^{(2)}\}|,$$

where $|\cdot|$ denotes size of a set. Note that $0 \leq R_i^{j,(2)} \leq n$ for $i = 1, 2, \dots, n, j = 1, 2$ and also that since $R_i^{1,(2)} \vee R_i^{2,(2)} = |\{j : X_j^{(2)} \geq X_i^{(2)}\}|$, $1 \leq R_i^{1,(2)} \vee R_i^{2,(2)} \leq n$ for $i = 1, 2, \dots, n$. Proposition 2.9 gives an estimator of $\tilde{\nu}^0$ which we modify to get an estimator of $S^0(\cdot)$.

Proposition 2.9. *We have in $M_+((0, \infty]^2)$,*

$$(2.25) \quad \widehat{\tilde{\nu}}_n^0(\cdot) := \frac{1}{k} \sum_{i=1}^n \epsilon_{(k/R_i^{1,(2)}, k/R_i^{2,(2)})}(\cdot) \Rightarrow \tilde{\nu}^0(\cdot) \quad (k \rightarrow \infty, n/k \rightarrow \infty),$$

where (2.12) defines $\tilde{\nu}^0(\cdot)$ and (2.24) defines $R_i^{1,(2)}$ and $R_i^{2,(2)}$.

Proof. From (2.5) and the definition of $\psi^0(\cdot)$ given in (2.7), we have in $D((0, \infty])$,

$$\frac{1}{k} \sum_{i=1}^n \epsilon_{((X_i^{(2)} - d(n/k))/c(n/k))}((x, \infty]) \Rightarrow [\psi^0(x)]^{-1}.$$

Hence [23, page 58], inverse functions also converge in distribution in $D_{\text{left}}((0, \infty])$, the space of left continuous functions with finite right limits,

$$(2.26) \quad \inf\{x : \frac{1}{k} \sum_{i=1}^n \epsilon_{((X_i^{(2)} - d(n/k))/c(n/k))}((x, \infty]) \leq 1/s\} \Rightarrow \inf\{x : [\psi^0(x)]^{-1} \leq 1/s\} = (\psi^0)^{\leftarrow}(s).$$

Write the order statistics of $\{X_1^{(2)}, \dots, X_n^{(2)}\}$ as $X_{(1)}^{(2)} \geq \dots \geq X_{(n)}^{(2)}$ and observe the left side of (2.26) is

$$(2.27) \quad \inf\{x : \sum_{i=1}^n \epsilon_{((X_i^{(2)} - d(n/k))/c(n/k))}((x, \infty]) \leq k/s\} = \left(\frac{X_{(\lceil k/s \rceil)}^{(2)} - d(n/k)}{c(n/k)} \right).$$

From (2.20), (2.26) and (2.27) we get that as $k \rightarrow \infty$ and $n/k \rightarrow \infty$,

$$(2.28) \quad \left(\frac{1}{k} \sum_{i=1}^n \epsilon_{\left(\frac{X_i^{(2)} - d(n/k)}{c(n/k)}\right)}(\cdot), \frac{X_{(\lceil k/s \rceil)}^{(2)} - d(n/k)}{c(n/k)}, \frac{X_{(\lceil k/t \rceil)}^{(2)} - d(n/k)}{c(n/k)} \right) \Rightarrow (\nu^0(\cdot), ((\psi^0)^{\leftarrow}(s), (\psi^0)^{\leftarrow}(t)))$$

in $M_+(\mathbb{E}^0) \times D_{\text{left}}((0, \infty]) \times D_{\text{left}}((0, \infty])$. Using the scaling technique as in [23, page 311] we get from (2.28) that as $k \rightarrow \infty$ and $n/k \rightarrow \infty$,

(2.29)

$$\frac{1}{k} \sum_{i=1}^n 1_{\{X_i^1 > X_{(\lceil k/s \rceil)}^{(2)}, X_i^2 > X_{(\lceil k/t \rceil)}^{(2)}\}} \Rightarrow \nu^0(((\psi^0)^\leftarrow(s), \infty] \times ((\psi^0)^\leftarrow(t), \infty]) = \tilde{\nu}^0((s, \infty] \times (t, \infty]),$$

in $D_{\text{left}}(0, \infty] \times D_{\text{left}}((0, \infty])$. Since the left side of (2.29) is

$$\frac{1}{k} \sum_{i=1}^n 1_{\{R_i^{1,(2)} < k/s, R_i^{2,(2)} < k/t\}} = \frac{1}{k} \sum_{i=1}^n 1_{\{s < k/R_i^{1,(2)}, t < k/R_i^{2,(2)}\}},$$

we have proven (2.25). \square

Proposition 2.9 yields an estimator of the limit measure $\nu_1 \times S^0(\cdot)$ and then an estimator of $S^0(\cdot)$.

Proposition 2.10. *The convergence in (2.25) is equivalent to*

$$(2.30) \quad \widehat{\nu_1 \times S^0}_n(\cdot) := \frac{1}{k} \sum_{i=1}^n \epsilon \left(\frac{k}{R_i^{1,(2)} \vee R_i^{2,(2)}}, \left(\frac{R_i^{1,(2)} \vee R_i^{2,(2)}}{R_i^{1,(2)}}, \frac{R_i^{1,(2)} \vee R_i^{2,(2)}}{R_i^{2,(2)}} \right) \right) (\cdot) \Rightarrow \nu_1 \times S^0(\cdot)$$

in $M_+((0, \infty] \times \delta\mathbb{N}^{(2)})$, where $\nu_1 \times S^0(\cdot)$ is given in Proposition 2.5.

Proof. The proof uses Proposition 2.5 and follows exactly similar steps as that of Proposition 3.7 of [19]. It is based on the map $\mathbf{x} \mapsto (x^{(2)}, \mathbf{x}/x^{(2)})$. \square

From the convergence in (2.30), we construct a consistent estimator of $S^0(\cdot)$:

$$(2.31) \quad \hat{S}_n^0(\cdot) := \frac{\sum_{i=1}^n \epsilon \left(\frac{k}{R_i^{1,(2)} \vee R_i^{2,(2)}}, \left(\frac{R_i^{1,(2)} \vee R_i^{2,(2)}}{R_i^{1,(2)}}, \frac{R_i^{1,(2)} \vee R_i^{2,(2)}}{R_i^{2,(2)}} \right) \right) ([1, \infty] \times \cdot)}{\sum_{i=1}^n \epsilon \frac{k}{R_i^{1,(2)} \vee R_i^{2,(2)}} ([1, \infty])} \Rightarrow S^0(\cdot)$$

in $M_+(\delta\mathbb{N}^{(2)})$. Hence, we have obtained a consistent estimator for both the extreme value index γ^0 and the standardized hidden spectral measure S^0 .

It is possible that $R_i^{j,(2)} = 0$ for some $j = 1, 2$ and some $i = 1, 2, \dots, n$ and thus division by zero may be indicated in (2.31). Though theoretically justified, this is not desirable when writing code for an estimator. The continuous bijection $T : \delta\mathbb{N}^{(2)} \mapsto [0, 1]$ given by $T : \mathbf{x} \mapsto x^2/(x^1 + x^2)$ provides an instant remedy. We use the convention that $\infty/\infty = 1$ and $1/\infty = 0$. Using this transformation, (2.31) becomes

$$(2.32) \quad \frac{\sum_{i=1}^n \epsilon \left(\frac{k}{R_i^{1,(2)} \vee R_i^{2,(2)}}, \left(\frac{R_i^{1,(2)}}{R_i^{1,(2)} + R_i^{2,(2)}} \right) \right) ([1, \infty] \times \cdot)}{\sum_{i=1}^n \epsilon \frac{k}{R_i^{1,(2)} \vee R_i^{2,(2)}} ([1, \infty])} \Rightarrow S^0 \circ T^{-1}(\cdot)$$

in $M_+([0, 1])$. Since T is a continuous bijection, we retrieve S^0 from $S^0 \circ T^{-1}$.

3. NON-STANDARD HIDDEN DOMAIN OF ATTRACTION

To provide more scope for applications, the non-standard case no longer assumes that $X^1 \stackrel{d}{=} X^2$. However, we have found that to construct a coherent estimation theory requires careful consideration of the definitions. As in the standard case, the goal is to approximate marginal and joint tail probabilities.

3.1. How to proceed? In order for (1.2) to hold when \mathbf{X} has different marginal distributions, one typically needs different centering and scaling constants for the two components of \mathbf{X} . Traditional theory [24, page 277, Proposition 5.15] proceeds by standardizing each component. However, a theory of hidden domain of attraction that follows this approach encounters problems in the estimation procedure that we could not resolve without strong second order conditions. Here is more detail on the traditional approach.

3.1.1. Trying the traditional approach. Suppose we try to proceed by the usual standardization procedure. As in the introduction, set $U^i(x) = 1/P[X^i > x]$ and define $\tilde{X}^i = U^i(X^i)$, $i = 1, 2$. The MEVT condition (1.2) is equivalent to standard regular variation on $\mathbb{E} = [\mathbf{0}, \infty] \setminus \{\mathbf{0}\}$; that is of vague convergence of $nP[(\tilde{\mathbf{X}}/n \in \cdot)]$ where $\tilde{\mathbf{X}} = (\tilde{X}^1, \tilde{X}^2)$ and the vague limit of $nP[(\tilde{\mathbf{X}}/n \in \cdot)]$ is a transformed version of the measure ν in (1.2); the transformation depends on the marginal domains of attraction in (1.2). See [10], [24, Chapter 5], [9]. This reduction to Pareto scale brings us to the setup of Section 2 and to define HDA we would add the analogue of equation (2.5) with $\tilde{\mathbf{X}}$ replacing \mathbf{X} . So far so good and this was the route followed for HRV in [11].

For estimation, the change to Pareto scale creates difficulties. Either one estimates U^i which introduces errors that are difficult to quantify, or one resorts to the non-parametric rank transform [9, 11, 14, 23] which requires a large data set and destroys independence. For estimation, the ranks method replaces the observed sample $\{\tilde{\mathbf{X}}_i, i = 1, \dots, n\}$ with $\{(1/r_i^1, 1/r_i^2), i = 1, \dots, n\}$ where $r_i^1 = |\{j : X_j^1 > X_i^1\}|$ with a similar definition for r_i^2 . This replacement creates the difficulty that it is not obvious the information about convergence in (2.5) on \mathbb{E}^0 is preserved. We have not found a bridge such as Proposition 1, page 401 in [11] and our efforts in this direction indicate a need for second order regular variation for consistency of estimates, an undesirable feature. So we try a different path that will be helpful for some cases.

3.1.2. An untraditional approach. We deviate from the traditional MEVT treatment by requiring that both components in (2.1) have the same centering and scaling but permitting the limit measure to have one zero marginal. By a zero marginal, we mean that either the limit measure $\nu(\cdot)$ in (2.1) has the property

$$(3.1) \quad \nu(\{z \in \mathbb{E} : z^2 > y\}) = 0 \quad ((y, y) \in \mathbb{E})$$

or the same holds with z^1 in place of z^2 . This could happen if the tail of X^2 is lighter than that of X^1 or vice versa. If $\mathbb{E} = [\mathbf{0}, \infty] \setminus \{\mathbf{0}\}$, (3.1) means ν must concentrate either on one of the two axes emanating from $\mathbf{0}$. The approach that we outline is not the only way to proceed but it does overcome some difficulties for estimation.

In the non-standard case, if we assume (2.1), the limit measure ν may satisfy:

- (i) ν has a zero second marginal: for $(y, y) \in \mathbb{E}$, $\nu(\{z \in \mathbb{E} : z^2 > y\}) = 0$;
- (ii) ν has a zero first marginal: for $(y, y) \in \mathbb{E}$, $\nu(\{z \in \mathbb{E} : z^1 > y\}) = 0$,
- (iii) the cases (i) and (ii) do not hold, but $\nu(\{z \in \mathbb{E} : z^1 > x, z^2 > y\}) = 0$ for $(x, y) \in \mathbb{E}$,
- (iv) for $(x, y) \in \mathbb{E}$, $\nu(\{z \in \mathbb{E} : z^1 > x, z^2 > y\}) > 0$.

Case (iv) means (2.1) yields non-zero estimates of the marginal and joint tail probabilities, so in this case we have no need to define HDA. The definition and analysis of HDA in case (iii) is the same as the standard case discussed in Section 2 since the two components are normalized the same. The definition and analysis of HDA are very similar for cases (i) and (ii) so focus only on case (i).

Our definition of HDA borrows a basic idea of the conditional extreme value (CEV) model [4–6, 12, 13]. A relevant state space is \mathbb{E}^\square where either $\mathbb{E}^\square = [-\infty, \infty] \times (-\infty, \infty]$ or $\mathbb{E}^\square =$

$[0, \infty] \times (0, \infty]$. We will see in Example 3.7 that for case (i) it is possible to find HDA on both the cones \mathbb{E}^\square and \mathbb{E}^0 in sequence. Thus, compared to the standard case, our estimation procedure here might involve analyzing HDA on the additional cone \mathbb{E}^\square .

Definition 3.1. The distribution of $\mathbf{X} = (X^1, X^2)$ has hidden domain of attraction on the cone \mathbb{E}^\square if (2.1) holds with the limit measure ν , the second marginal of ν is a zero measure and in addition, there exist constants $e_n > 0$ and $f_n \in \mathbb{R}$ and a non-zero measure $\nu^\square \in M_+(\mathbb{E}^\square)$ such that as $n \rightarrow \infty$,

$$(3.2) \quad nP \left[\left(\frac{\mathbf{X} - f_n \mathbf{1}}{e_n} \right) \in \cdot \right] \xrightarrow{v} \nu^\square(\cdot) \quad \text{in } M_+(\mathbb{E}^\square).$$

Note (3.2) does not preclude a similar convergence with different norming constants from holding on a smaller cone, say \mathbb{E}^0 . See Definition 3.3. From (3.2) it follows that for $(y, y) \in \mathbb{E}^\square$, as $n \rightarrow \infty$,

$$(3.3) \quad nP[X^2 > e_n y + f_n] \rightarrow \nu^\square(\{(u, v) \in E^\square : v > y\}).$$

Therefore, the distribution of X^2 , the second component of \mathbf{X} , belongs to the maximal domain of attraction of an extreme value distribution [9, page 4]. Using one-dimensional extreme value theory, $\psi^\square(y) := [\nu^\square([0, \infty] \times (y, \infty))]^{-1}$ must take one of the following forms [24, page 9]:

$$(3.4) \quad \begin{aligned} \psi^\square(y) &= \begin{cases} y^{1/\gamma^\square}, & \text{if } y > 0, \\ 0, & \text{otherwise,} \end{cases} & \text{if } \gamma^\square > 0, \\ \psi^\square(y) &= e^y, \ y \in \mathbb{R}, & \text{if } \gamma^\square = 0, \\ \psi^\square(y) &= \begin{cases} \infty, & \text{if } y > 0, \\ (-y)^{1/\gamma^\square}, & \text{otherwise,} \end{cases} & \text{if } \gamma^\square < 0. \end{aligned}$$

The parameter γ^\square in (3.4) is the extreme value index of X^2 . We can and always do choose $\{e_n\}$ and $\{f_n\}$ in such a way that ψ^\square takes one of the above forms.

Remark 3.2. Some comments about Definition 3.1.

- (1) Definition 3.1 often permits non-zero approximation of joint tail probabilities when asymptotic independence holds in (1.2) with each component normalized differently.
- (2) Since (2.1) holds, so does (2.3) and therefore the maximum component $X^{(1)}$ belongs to the maximal domain of attraction of some extreme value distribution. Since $X^{(1)} \geq X^2$, the convergence relation (2.1) constrains the possible convergences in (3.2). For example, if (2.3) has $X^{(1)}$ in the Gumbel domain of attraction, then the distribution of X^2 cannot have a regularly varying tail.
- (3) The distribution of X^2 determines the cone \mathbb{E}^\square and (3.3) yields the scaling and centering constants $\{e_n\}$ and $\{f_n\}$. If the distribution of X^2 is in the Fréchet domain of attraction, $\mathbb{E}^\square = [0, \infty] \times (0, \infty]$ and otherwise we take $\mathbb{E}^\square = [-\infty, \infty] \times (-\infty, \infty]$.

There are two possibilities for the limit measure ν^\square in (3.2):

- (i) the limit measure ν^\square puts zero mass on all sets $(x, \infty] \times (y, \infty]$ for $(x, y) \in \mathbb{E}^\square$; or
- (ii) the limit measure ν^\square puts non-zero mass on one of the sets $(x, \infty] \times (y, \infty]$ for $(x, y) \in \mathbb{E}^\square$.

The semi-parametric structure of ν^\square discussed in the next section implies that for case (ii), $\nu^\square((x, \infty] \times (y, \infty]) > 0$ for all $(x, y) \in \mathbb{E}^\square$. So, in case (ii), we get non-zero estimates of joint tail probabilities and since we accomplished our goal there is no reason to seek further instances of HDA. Case (i) offers the difficulty that $\nu^\square(\cdot)$ fails to provide non-zero estimates of joint tail probabilities. However, HDA could still exist on a smaller cone such as \mathbb{E}^0 and if this is true we have a potential resolution of the difficulty. We formalize this idea in Definition 3.3, where the state space \mathbb{E}^0 is either $(-\infty, \infty]^2$ or $(0, \infty]^2$.

Definition 3.3. The distribution of $\mathbf{X} = (X^1, X^2)$ has hidden domain of attraction on the cones \mathbb{E}^\square and \mathbb{E}^0 if Definition 3.1 holds, the limit measure ν^\square in (3.2) puts zero mass on sets of the form $(x, \infty] \times (y, \infty]$ for $(x, y) \in \mathbb{E}^\square$, and in addition, there exist centering and scaling constants $\{c_n\}$ and $\{d_n\}$ and a non-zero measure $\nu^0 \in M_+(\mathbb{E}^0)$ such that as $n \rightarrow \infty$,

$$(3.5) \quad nP \left[c_n^{-1} (\mathbf{X} - d_n \mathbf{1}) \in \cdot \right] \xrightarrow{v} \nu^0(\cdot) \quad \text{in } M_+(\mathbb{E}^0).$$

As noted before in (2.6), the scaling and centering constants $\{c_n\}$ and $\{d_n\}$ in (3.5) are characterized by the distribution of $X^{(2)}$, the minimum component of \mathbf{X} . Recall the definition of ψ^0 given in (2.7). As was done in the standard case discussion, we choose the scaling and centering constants $\{c_n\}$ and $\{d_n\}$ in (3.5) so that ψ^0 takes one of the forms given in (2.9). Also, whether \mathbb{E}^0 in (3.5) is $(-\infty, \infty]^2$ or $(0, \infty]^2$ is determined by the distribution of $X^{(2)}$.

3.2. Semi-parametric structure of ν^\square . Both limit measures ν^\square of (3.2) and ν^0 of (3.5) have semi-parametric structures. Since the semi-parametric structure of ν^0 was discussed in Section 2.1, we concentrate only on the semi-parametric structure of ν^\square and proceed as follows.

Recall that the distributions of \mathbf{X} and X^2 are H and H^2 . Define

$$(3.6) \quad U^2(x) = 1/(1 - H^2(x)).$$

The following proposition relates (3.2) to a regular variation condition on $[0, \infty] \times (0, \infty]$. Its proof is similar to that of Proposition 2.3 and is omitted.

Proposition 3.4. *The distribution of \mathbf{X} satisfies (3.2) iff*

- (1) X^2 has a distribution in a maximal domain of attraction so that for some $e_n > 0$, $f_n \in \mathbb{R}$ (3.3) holds, and
- (2) regular variation on the cone $[0, \infty] \times (0, \infty]$ holds for the distribution of $(U^2(X^1), U^2(X^2))$:

$$(3.7) \quad nP \left[\left(\frac{U^2(X^1)}{n}, \frac{U^2(X^2)}{n} \right) \in \cdot \right] \xrightarrow{v} \tilde{\nu}^\square(\cdot) \quad (\text{in } M_+([0, \infty] \times (0, \infty])),$$

where (3.6) defines $U^2(\cdot)$ and $\tilde{\nu}^\square(\cdot)$ is a Radon measure on $[0, \infty] \times (0, \infty]$.

The limit measure $\tilde{\nu}^\square(\cdot)$ is related to the limit measure in $\nu^\square(\cdot)$ in (3.2) by the following relation: for $(x^1, x^2) \in [0, \infty] \times (0, \infty]$,

$$(3.8) \quad \begin{aligned} \tilde{\nu}^\square((x^1, \infty] \times (x^2, \infty]) &= \nu^\square \left(((\psi^\square)^\leftarrow(x^1), \infty] \times ((\psi^\square)^\leftarrow(x^2), \infty] \right), \\ \tilde{\nu}^\square([0, x^1] \times (x^2, \infty]) &= \nu^\square \left(\{z \in \mathbb{E}^\square : z^1 \leq (\psi^\square)^\leftarrow(x^1), z^2 > (\psi^\square)^\leftarrow(x^2)\} \right). \end{aligned}$$

The measure $\tilde{\nu}^\square(\cdot)$ satisfies the scaling property:

$$(3.9) \quad \tilde{\nu}^\square(c \cdot) = c^{-1} \tilde{\nu}^\square(\cdot) \quad c > 0.$$

Remark 3.5. (i) On the semi-parametric structure of ν^\square : We will see that a probability measure S^\square on $[0, \infty]$ determines the limit measure $\tilde{\nu}^\square$. The parameter γ^\square and the probability measure S^\square on $[0, \infty]$ determine ν^\square , since given γ^\square and measure S^\square , we get the function $\psi^\square(\cdot)$ in (3.4) and $\tilde{\nu}^\square$ which through (3.8) determines ν^\square .

- (ii) If the support of the distribution of X^2 is smaller than that of X^1 , then $U^2(X^1)$ could take the value ∞ , so in this case we consider $U^2(X^1)$ as an extended random variable.

The method that shows $\tilde{\nu}^0([1, \infty]^2) = 1$ also shows $\tilde{\nu}^\square([0, \infty] \times [1, \infty]) = 1$. Proposition 4 of [12] and Proposition 3.4 give a convergence relation in new coordinates.

Proposition 3.6. *The convergence in (3.2) is equivalent to*

$$(3.10) \quad nP \left[\left(\frac{U^2(X^2)}{n}, \frac{U^2(X^1)}{U^2(X^2)} \right) \in \cdot \right] \xrightarrow{v} \nu_1 \times S^\square(\cdot) \quad (\text{in } M_+((0, \infty] \times [0, \infty])),$$

where ν_1 is a Pareto measure on $(0, \infty]$ satisfying $\nu_1((x, \infty]) = x^{-1}$ for $x > 0$, and S^\square is a probability measure on $[0, \infty]$, called the standardized hidden spectral measure. The relation between $\tilde{\nu}^\square$ given in (2.12) and S^\square is

$$(3.11) \quad \tilde{\nu}^\square(\{\mathbf{x} \in [0, \infty] \times (0, \infty] : x^2 \geq r, x^1/x^2 \in \Lambda\}) = r^{-1} S^\square(\Lambda),$$

which holds for all $r > 0$ and all Borel sets $\Lambda \subset [0, \infty]$.

3.3. Examples. We give examples of distributions of $\mathbf{X} = (X^1, X^2)$, $X^1 \stackrel{d}{\neq} X^2$, and which have HDA. In Example 3.7, the limit measure ν^\square of (3.2) puts zero mass on all sets of the form $(x^1, \infty] \times (x^2, \infty]$ for $\mathbf{x} \in \mathbb{E}^\square$ and HDA also holds on \mathbb{E}^0 . In Example 3.8, ν^\square of (3.2) puts non-zero mass on sets of the form $(x^1, \infty] \times (x^2, \infty]$ for $\mathbf{x} \in \mathbb{E}^\square$.

Example 3.7. Suppose $X^1 \sim \exp(1)$, $X^2 \sim \exp(2)$ and (X^1, X^2) are independent. Hence (X^1, X^2) are also asymptotically independent and to estimate joint tail probabilities, we proceed as follows. For $\mathbf{x} \in \mathbb{R}^2$,

$$\begin{aligned} n(1 - P[X^1 - \log n \leq x^1, X^2 - \log n \leq x^2]) &= n[1 - (1 - e^{-(\log n + x^1)})(1 - e^{-2(\log n + x^2)})] \\ &\sim n(e^{-(\log n + x^1)} + e^{-2(\log n + x^2)}) \rightarrow e^{-x^1} \quad \text{as } n \rightarrow \infty, \end{aligned}$$

which implies (2.1) holds on $\mathbb{E} = [-\infty, \infty]^2 \setminus \{(-\infty, -\infty)\}$ with $\nu(([-\infty, x^1] \times [-\infty, x^2])^c) = e^{-x^1}$ and ν puts mass only on $(-\infty, \infty] \times \{-\infty\}$, the horizontal line through $-\infty$, and ν has zero second marginal. So we seek HDA on \mathbb{E}^\square . For $\frac{1}{2} \log n + x^i > 0$, $i = 1, 2$ we have

$$nP \left[X^1 - \frac{\log n}{2} \leq x^1, X^2 - \frac{\log n}{2} > x^2 \right] = n \left(1 - e^{-(\frac{\log n}{2} + x^1)} \right) e^{-2(\frac{\log n}{2} + x^2)} \rightarrow e^{-2x^2},$$

as $n \rightarrow \infty$, and

$$nP \left[X^1 - \frac{\log n}{2} > x^1, X^2 - \frac{\log n}{2} > x^2 \right] = n e^{-(\frac{\log n}{2} + x^1)} e^{-2(\frac{\log n}{2} + x^2)} \rightarrow 0.$$

Thus, HDA exists on $\mathbb{E}^\square = [-\infty, \infty] \times (-\infty, \infty]$ with limit measure ν^\square , where $\nu^\square([- \infty, x^1] \times (x^2, \infty]) = e^{-2x^2}$ and $\nu^\square((x^1, \infty] \times (x^2, \infty]) = 0$ for $\mathbf{x} \in \mathbb{E}^\square$. So ν^\square concentrates on $\{-\infty\} \times (-\infty, \infty]$, the vertical line through $-\infty$. After peeling away both lines through $-\infty$, we look for HDA on \mathbb{E}^0 . A hint for how to proceed is provided by $X^1 \wedge X^2 \sim \exp(3)$. Note that as $n \rightarrow \infty$,

$$nP \left[X^1 - \frac{\log n}{3} > x^1, X^2 - \frac{\log n}{3} > x^2 \right] = n e^{-(\frac{\log n}{3} + x^1)} e^{-2(\frac{\log n}{3} + x^2)} \rightarrow e^{-(x^1 + 2x^2)}.$$

Thus HDA exists on $\mathbb{E}^0 = (-\infty, \infty]^2$ with limit measure ν^0 , where $\nu^0((x^1, \infty] \times (x^2, \infty]) = e^{-(x^1 + 2x^2)}$ for $\mathbf{x} \in \mathbb{E}^0$.

To summarize Example 3.7, Definition 3.3 holds and HDA holds on both the cones \mathbb{E}^\square and \mathbb{E}^0 , but the HDA on \mathbb{E}^\square is not informative for calculating risk probabilities where both components of the risk vector are large.

Example 3.8. Suppose E_1, E_2, E_3 are iid $\exp(1)$ random variables independent of $B \sim \text{Bernoulli}(1/2)$ and define \mathbf{X} as

$$\mathbf{X} = B(E_1, E_3/3) + (1 - B)(E_2/2, E_2/2).$$

Then \mathbf{X} possesses asymptotic independence and to estimate joint tail probabilities, proceed as follows. As $n \rightarrow \infty$, for $x^1 \wedge x^2 + \log n > 0$,

$$\begin{aligned} n(1 - P[X^1 - \log n \leq x^1, X^2 - \log n \leq x^2]) \\ = n \left[1 - \left[\frac{1}{2} \left(1 - e^{-(\log n + x^1)} \right) \left(1 - e^{-3(\log n + x^2)} \right) + \frac{1}{2} \left(1 - e^{-2(\log n + x^1 \wedge x^2)} \right) \right] \right] \rightarrow \frac{1}{2} e^{-x^1}, \end{aligned}$$

which implies (2.1) holds on $\mathbb{E} = [-\infty, \infty]^2 \setminus \{(-\infty, -\infty)\}$ with $\nu(([-\infty, x^1] \times [-\infty, x^2])^c) = \frac{1}{2} e^{-x^1}$ and ν concentrates on $(-\infty, \infty] \times \{-\infty\}$. Thus ν has zero second marginal and we seek HDA on \mathbb{E}^\square . Note as $n \rightarrow \infty$,

$$\begin{aligned} nP \left[X^1 - \frac{\log n}{2} \leq x^1, X^2 - \frac{\log n}{2} > x^2 \right] \\ = n \left[\frac{1}{2} \left(1 - e^{-(\frac{\log n}{2} + x^1)} \right) e^{-3(\frac{\log n}{2} + x^2)} + \frac{1}{2} \left(e^{-2(\frac{\log n}{2} + x^2)} - e^{-2(\frac{\log n}{2} + x^1)} \right) 1_{\{x^2 < x^1\}} \right] \\ \rightarrow \frac{1}{2} \left(e^{-2x^2} - e^{-2x^1} \right) 1_{\{x^2 < x^1\}} \end{aligned}$$

and therefore, $nP[X^1 - \frac{1}{2} \log n > x^2] \rightarrow \frac{1}{2} \exp\{-2x^2\}$ and consequently,

$$nP \left[X^1 - \frac{\log n}{2} > x^1, X^2 - \frac{\log n}{2} > x^2 \right] \rightarrow \frac{1}{2} e^{-2(x^1 \vee x^2)}.$$

Thus, HDA exists on $\mathbb{E}^\square = [-\infty, \infty] \times (-\infty, \infty]$ with limit measure ν^\square , where $\nu^\square([-\infty, x^1] \times (x^2, \infty]) = \frac{1}{2} (e^{-2x^2} - e^{-2x^1}) 1_{\{x^2 < x^1\}}$ and $\nu^\square((x^1, \infty] \times (x^2, \infty]) = \frac{1}{2} e^{-2(x^1 \vee x^2)}$ for $\mathbf{x} \in \mathbb{E}^\square$. In fact, ν^\square concentrates on the line $\{(x, x) : x \in (-\infty, \infty)\}$.

Since $\nu^\square((x^1, \infty] \times (x^2, \infty]) > 0$ for $\mathbf{x} \in \mathbb{E}^\square$, we do not seek HDA on \mathbb{E}^0 .

3.4. Estimation methods. To estimate joint tail probabilities, we require an estimate of the limit measure ν^\square given in Definition 3.1 and possibly ν^0 given in Definition 3.3. Estimation of ν^0 follows the same steps as in Section 2.3 so we concentrate on estimating ν^\square . Let, $\{\mathbf{X}, \mathbf{X}_i, i = 1, 2, \dots, n\}$ be iid where the distribution of \mathbf{X} satisfies (3.2). From (3.2) we get [23, page 139] in $M_+(\mathbb{E}^\square)$

$$(3.12) \quad \frac{1}{k} \sum_{i=1}^n \epsilon \left(\frac{X_i^1 - f(n/k)}{e(n/k)}, \frac{X_i^2 - f(n/k)}{e(n/k)} \right) (\cdot) \Rightarrow \nu^\square(\cdot) \quad (k \rightarrow \infty, n/k \rightarrow \infty).$$

We know from (3.3) that the distribution of X^2 characterizes $\{e_n\}$ and $\{f_n\}$ and from the iid data $\{X_i^2 : i = 1, 2, \dots, n\}$, we can construct estimators of $e(n/k)$ and $f(n/k)$ denoted by $\hat{e}(n/k)$ and $\hat{f}(n/k)$ [23, page 93] such that

$$(3.13) \quad \frac{e(n/k)}{\hat{e}(n/k)} \xrightarrow{P} 1, \quad \frac{f(n/k) - \hat{f}(n/k)}{e(n/k)} \xrightarrow{P} 0.$$

Since the limits in (3.13) are constants, we get joint convergence in $M_+(\mathbb{E}^\square) \times \mathbb{R}^2$,

$$(3.14) \quad \left(\frac{1}{k} \sum_{i=1}^n \epsilon \left(\frac{X_i^1 - f(n/k)}{e(n/k)}, \frac{X_i^2 - f(n/k)}{e(n/k)} \right), \frac{f(n/k) - \hat{f}(n/k)}{e(n/k)}, \frac{e(n/k)}{\hat{e}(n/k)} \right) \Rightarrow (\nu^\square(\cdot), 0, 1)$$

Apply the continuous mapping theorem to (3.14) using the map $(\nu(\cdot), b, a) \mapsto \nu(a[(\cdot) + b])$ to get in $M_+(\mathbb{E}^\square)$

$$(3.15) \quad \widehat{\nu}_n^\square(\cdot) := \frac{1}{k} \sum_{i=1}^n \epsilon\left(\frac{X_i^1 - \bar{f}(n/k)}{\bar{e}(n/k)}, \frac{X_i^2 - \bar{f}(n/k)}{\bar{e}(n/k)}\right)(\cdot) \Rightarrow \nu^\square(\cdot) \quad (k \rightarrow \infty, n/k \rightarrow \infty).$$

This estimator of ν^\square is non-parametric and as in Section 2.3, we exploit the semi-parametric structure of ν^\square by estimating γ^\square and the standardized hidden spectral measure S^\square . The parameter γ^\square is the extreme value index of the distribution of X^2 so estimating this from iid data $\{X_i^2 : i = 1, 2, \dots, n\}$ is standard [9, page 65]. We obtain an estimator of $S^\square(\cdot)$ by modifying (2.24) to account for the difference between \mathbb{E}^0 and \mathbb{E}^\square . Since the first step is to construct a consistent estimator of $\tilde{\nu}^\square(\cdot)$ defined in (3.7), define

$$(3.16) \quad R_i^{1,2} := \left| \{j : X_j^2 \geq X_i^1\} \right| \quad \text{and} \quad R_i^{2,2} := \left| \{j : X_j^2 \geq X_i^2\} \right|, \quad (i = 1, 2, \dots, n,)$$

where $|\cdot|$ denotes size of a set. Observe $R_i^{2,2}$ is just the anti-rank of X_i^2 and thus $1 \leq R_i^{2,2} \leq n$ for $i = 1, 2, \dots, n$. Also, $0 \leq R_i^{1,2} \leq n$ for $i = 1, 2, \dots, n$. An estimator of $\tilde{\nu}^\square$ is obtained from the convergence in $M_+([0, \infty] \times (0, \infty])$:

$$(3.17) \quad \widehat{\tilde{\nu}}_n^\square(\cdot) := \frac{1}{k} \sum_{i=1}^n \epsilon_{(k/R_i^{1,2}, k/R_i^{2,2})}(\cdot) \Rightarrow \tilde{\nu}^\square(\cdot) \quad (k \rightarrow \infty, n/k \rightarrow \infty).$$

The verification of (3.17) follows the steps used in the proof of Proposition 2.9. Changing coordinate system in (3.17) leads to an estimator of $\nu_1 \times S^\square(\cdot)$ from the convergence in $M_+((0, \infty] \times [0, \infty])$

$$(3.18) \quad \widehat{\nu_1 \times S^\square}_n(\cdot) := \frac{1}{k} \sum_{i=1}^n \epsilon_{(k/R_i^{2,2}, R_i^{2,2}/R_i^{1,2})}(\cdot) \Rightarrow \nu_1 \times S^\square(\cdot) \quad (k \rightarrow \infty, n/k \rightarrow \infty)$$

and this produces an estimator of S^\square since in $M_+([0, \infty])$,

$$(3.19) \quad \widehat{S}_n^\square(\cdot) := \frac{\sum_{i=1}^n \epsilon_{(k/R_i^{2,2}, R_i^{2,2}/R_i^{1,2})}([1, \infty] \times \cdot)}{\sum_{i=1}^n \epsilon_{k/R_i^{2,2}}([1, \infty])} \Rightarrow S^\square(\cdot) \quad (k \rightarrow \infty, n/k \rightarrow \infty).$$

This estimator may be modified as in (2.32) using the continuous bijection $TR : [0, \infty] \mapsto [0, 1]$ defined by $TR : x \mapsto x/(1+x)$ to get in $M_+([0, 1])$,

$$(3.20) \quad \widehat{S}_n^\square(\cdot) \circ TR^{-1} \Rightarrow S^\square \circ TR^{-1}(\cdot).$$

This summarizes how to obtain consistent estimators for extreme value index γ^\square and the standardized hidden spectral measure S^\square .

4. DETECTION OF HDA

Since HDA is a generalization of HRV, it is not surprising that the detection techniques have similarities to those used for HRV; see [19] and [23, pages 316-340]. However, we deviate from traditional MEVT and so we proceed carefully. Instead of (1.2), we are assuming (2.1) and allowing for the possibility that the limit measure ν has a zero marginal.

A first step to using HDA to compute joint tail probabilities is to infer where the limit measure ν in (2.1) concentrates mass. When ν concentrates on one or more lines, we seek HDA. Traditionally the inference about the support of ν has often been done informally with a density plot of a spectral measure after non-parametric transformation to Pareto scale [23, pages 316-321].

We consider an appropriate spectral measure for the task of understanding the support as follows. Since (2.1) implies $X^{(1)}$ has a distribution $H^{(1)}$ in a maximal domain of attraction, we define $U^{(1)}(\cdot) = 1/(1 - H^{(1)}(\cdot))$. From (2.1) we get on $(0, \infty]^2$ that

$$(4.1) \quad nP \left[\left(U^{(1)}(X^1)/n, U^{(1)}(X^2)/n \right) \in \cdot \right] \xrightarrow{v} \tilde{\nu}(\cdot),$$

where $\tilde{\nu}(\cdot)$ is a Radon measure on $(0, \infty]^2$ related to the limit measure $\nu(\cdot)$ in (2.1) by

$$(4.2) \quad \tilde{\nu}((x^1, \infty] \times (x^2, \infty]) = \nu((\psi^{\leftarrow}(x^1), \infty] \times (\psi^{\leftarrow}(x^2), \infty]) \quad (\mathbf{x} \in (0, \infty]^2),$$

and ψ is defined in (2.4) as $\psi(y) := (\nu(\{\mathbf{z} \in \mathbb{E} : z^{(1)} > y\}))^{-1}$. The measure $\tilde{\nu}(\cdot)$ satisfies the scaling

$$(4.3) \quad \tilde{\nu}(c \cdot) = c^{-1} \tilde{\nu}(\cdot), \quad c > 0,$$

and convergence in (4.1) is equivalent to

$$(4.4) \quad nP \left[\left(\frac{U^{(1)}(X^{(1)})}{n}, \left(\frac{U^{(1)}(X^1)}{U^{(1)}(X^{(1)})}, \frac{U^{(1)}(X^2)}{U^{(1)}(X^{(1)})} \right) \right) \in \cdot \right] \xrightarrow{v} \nu_1 \times S(\cdot),$$

on $(0, \infty] \times \delta\mathbb{N}^{(1)}$, where $\nu_1((y, \infty]) = y^{-1}$ for $y > 0$, $\delta\mathbb{N}^{(1)} = \{\mathbf{x} \in (0, \infty]^2 : x^{(1)} = 1\}$ and S is a probability measure on $\delta\mathbb{N}^{(1)}$. The standardized spectral measure S is related to $\tilde{\nu}$ in (4.1) by

$$(4.5) \quad \tilde{\nu} \left(\left\{ \mathbf{x} \in (0, \infty]^2 : x^{(1)} \geq r, \mathbf{x}/x^{(1)} \in \Lambda \right\} \right) = r^{-1} S(\Lambda), \quad r > 0, \text{ Borel set } \Lambda \subset \delta\mathbb{N}^{(1)}.$$

To estimate this measure $S(\cdot)$, we define variants of the anti-ranks

$$(4.6) \quad R_i^{1,(1)} = |\{j : X_j^{(1)} \geq X_i^1\}| \quad \text{and} \quad R_i^{2,(1)} = |\{j : X_j^{(1)} \geq X_i^2\}| \quad (1 \leq i \leq n).$$

A consistent estimator of $S(\cdot)$ is obtained from the convergence in $M_+(\delta\mathbb{N}^{(1)})$

$$(4.7) \quad \hat{S}_n := \frac{\sum_{i=1}^n \epsilon \left(\frac{k}{R_i^{1,(1)} \wedge R_i^{2,(1)}}, \left(\frac{R_i^{1,(1)} \wedge R_i^{2,(1)}}{R_i^{1,(1)}}, \frac{R_i^{1,(1)} \wedge R_i^{2,(1)}}{R_i^{2,(1)}} \right) \right) ([1, \infty] \times \cdot)}{\sum_{i=1}^n \epsilon \frac{k}{R_i^{1,(1)} \wedge R_i^{2,(1)}} ([1, \infty])} \Rightarrow S(\cdot),$$

($k \rightarrow \infty, n/k \rightarrow \infty$). The continuous bijection $T : \delta\mathbb{N}^{(1)} \mapsto [0, 1]$ given by $T : \mathbf{x} \mapsto x^2/(x^1 + x^2)$ transforms (4.7) to $\hat{S}_n \circ T^{-1} \Rightarrow S \circ T^{-1}(\cdot)$ in $M_+([0, 1])$. A density plot of $\hat{S}_n \circ T^{-1}$ is easier to analyze because $[0, 1]$ is a nicer space than $\delta\mathbb{N}^{(1)}$.

Analyzing the density plot using the points of $\hat{S}_n \circ T^{-1}$ should yield evidence falling into the following categories:

- (i) The distribution $S \circ T^{-1}$ concentrates near 0, so remove $\{(x, y) \in \mathbb{R}^2 : y = -\infty\}$ and seek HDA on $\mathbb{E}^\square = [-\infty, \infty] \times (-\infty, \infty]$ or its first quadrant analogue, depending on the distribution of the second component of the random vector; see Remark 3.2 (3).
- (ii) The distribution $S \circ T^{-1}$ concentrates near 1, so remove $\{(x, y) \in \mathbb{R}^2 : x = -\infty\}$ and seek HDA on $(-\infty, \infty] \times [-\infty, \infty]$ or its first quadrant analogue depending on the distribution of the first component of the random vector.
- (iii) The distribution $S \circ T^{-1}$ concentrates near 0 and 1, so remove $\{(x, y) \in \mathbb{R}^2 : x = -\infty\} \cup \{(x, y) \in \mathbb{R}^2 : y = -\infty\}$ and seek HDA on $\mathbb{E}^0 = (-\infty, \infty] \times (-\infty, \infty]$ or its first quadrant analogue, depending on the distribution of the smallest component of the random vector; see Remark 2.2 (2).
- (iv) The distribution $S \circ T^{-1}$ does not have any of the above properties; we have no evidence for the support of ν being restricted to at most 2 lines and we do not consider HDA.

We summarize our detection strategy and for concreteness assume $S \circ T^{-1}$ satisfies category (i): Check whether X^2 belongs to some maximal domain of attraction using a Hill or Pickands plots. If so, conclude HDA exists on \mathbb{E}^\square . Then consider whether HDA exists also on \mathbb{E}^0 by examining a similar kernel density plot formed by using the points of the estimator of $S^\square \circ (TR)^{-1}$ given in (3.20).

5. CONCLUSION

We defined HDA as a generalization of HRV and showed by example that for some random vectors, HDA exists but HRV does not. We outlined detection and estimation methods for HDA that show what is possible and emphasize that such methods fill a gap to provide improved estimates of probability of simultaneous exceedance by components of a risk vector. We have not implemented the methods nor demonstrated utility by analyzing data; this is a future project.

We restricted discussion to two dimensions and as observed for HRV [19], extensions to higher dimensions are not always straightforward and involve subtleties. In particular, in higher dimensions there are many more ways domains of attraction could be hidden and many more subspaces to explore for behavior that helps to estimate risk probabilities.

As with HRV [19, 20], our detection and estimation methods are exploratory and our estimators are only provably consistent. More formal statistical theory is needed to turn exploratory methods into confirmatory ones.

REFERENCES

- [1] A. A. Balkema and L. de Haan. On R. von Mises' condition for the domain of attraction of $\exp(-e^{-x})^1$. *Ann. Math. Statist.*, 43:1352–1354, 1972. ISSN 0003-4851.
- [2] P. Billingsley. *Convergence of Probability Measures*. John Wiley & Sons Inc., New York, second edition, 1999. ISBN 0-471-19745-9. A Wiley-Interscience Publication.
- [3] J.T. Bruun and J.A. Tawn. Comparison of approaches for estimating the probability of coastal flooding. *J. R. Stat. Soc., Ser. C, Appl. Stat.*, 47(3):405–423, 1998.
- [4] B. Das and S.I. Resnick. Conditioning on an extreme component: Model consistency with regular variation on cones. *Bernoulli*, 17(1):226–252, 2011. ISSN 1350-7265. doi: 10.3150/10-BEJ271.
- [5] B. Das and S.I. Resnick. Detecting a conditional extreme value model. *Extremes*, 14(1):29–61, 2011.
- [6] B. Das, A. Mitra, and S. Resnick. Living on the multi-dimensional edge: Seeking hidden risks using regular variation. *ArXiv e-prints 1108.5560*, August 2011.
- [7] L. de Haan. A characterization of multidimensional extreme-value distributions. *Sankhyā Ser. A*, 40(1):85–88, 1978. ISSN 0581-572X.
- [8] L. de Haan and J. de Ronde. Sea and wind: multivariate extremes at work. *Extremes*, 1(1):7–46, 1998.
- [9] L. de Haan and A. Ferreira. *Extreme Value Theory: An Introduction*. Springer-Verlag, New York, 2006.
- [10] L. de Haan and S.I. Resnick. Limit theory for multivariate sample extremes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 40(4):317–337, 1977.
- [11] J.E. Heffernan and S.I. Resnick. Hidden regular variation and the rank transform. *Adv. Appl. Prob.*, 37(2):393–414, 2005.
- [12] J.E. Heffernan and S.I. Resnick. Limit laws for random vectors with an extreme component. *Ann. Appl. Probab.*, 17(2):537–571, 2007. ISSN 1050-5164. doi: 10.1214/105051606000000835.

- [13] J.E. Heffernan and J.A. Tawn. A conditional approach for multivariate extreme values (with discussion). *JRSS B*, 66(3):497–546, 2004.
- [14] Xin Huang. *Statistics of Bivariate Extreme Values*. Ph.D. thesis, Tinbergen Institute Research Series 22, Erasmus University Rotterdam, Postbus 1735, 3000DR, Rotterdam, The Netherlands, 1992.
- [15] A.W. Ledford and J.A. Tawn. Statistics for near independence in multivariate extreme values. *Biometrika*, 83(1):169–187, 1996. ISSN 0006-3444.
- [16] A.W. Ledford and J.A. Tawn. Modelling dependence within joint tail regions. *J. Roy. Statist. Soc. Ser. B*, 59(2):475–499, 1997. ISSN 0035-9246.
- [17] A.W. Ledford and J.A. Tawn. Concomitant tail behaviour for extremes. *Adv. in Appl. Probab.*, 30(1):197–215, 1998. ISSN 0001-8678.
- [18] K. Maulik and S.I. Resnick. Characterizations and examples of hidden regular variation. *Extremes*, 7(1):31–67, 2005.
- [19] A. Mitra and S.I. Resnick. Hidden Regular Variation: Detection and Estimation. *Arxiv preprint arXiv:1001.5058*, 2010.
- [20] A. Mitra and S.I. Resnick. Hidden regular variation and detection of hidden risks. *Stochastic Models*, 27(4):591–614, 2011. doi: 10.1080/15326349.2011.614183. URL <http://www.tandfonline.com/doi/abs/10.1080/15326349.2011.614183>.
- [21] S.-H Poon, M. Rockinger, and J. Tawn. Modelling extreme-value dependence in international stock markets. *Statist. Sinica*, 13(4):929–953, 2003. ISSN 1017-0405. Statistical applications in financial econometrics.
- [22] S.I. Resnick. Hidden regular variation, second order regular variation and asymptotic independence. *Extremes*, 5(4):303–336 (2003), 2002. ISSN 1386-1999.
- [23] S.I. Resnick. *Heavy Tail Phenomena: Probabilistic and Statistical Modeling*. Springer Series in Operations Research and Financial Engineering. Springer-Verlag, New York, 2007. ISBN: 0-387-24272-4.
- [24] S.I. Resnick. *Extreme Values, Regular Variation and Point Processes*. Springer, New York, 2008. ISBN 978-0-387-75952-4. Reprint of the 1987 original.
- [25] R.L. Smith. Statistics of extremes, with applications in environment, insurance and finance. In B. Finkenstadt and H. Rootzén, editors, *SemStat: Seminaire Europeen de Statistique, Extreme Values in Finance, Telecommunications, and the Environment*, pages 1–78. Chapman-Hall, London, 2003.

ABHIMANYU MITRA, SCHOOL OF ORIE, CORNELL UNIVERSITY, ITHACA, NY 14853
E-mail address: am492@cornell.edu

SIDNEY I. RESNICK, SCHOOL OF ORIE, CORNELL UNIVERSITY, ITHACA, NY 14853
E-mail address: sir1@cornell.edu